



## WHITE PAPER

### How AI is Transforming eDiscovery

*Summary of Roundtable, hosted by ATARC on February 9, 2022*

During a recent roundtable discussion hosted by the Advanced Technology Academic Research Center (ATARC), participants from various federal agencies discussed current technological challenges associated with fully leveraging existing and available audio and video (AV) evidence in solving crimes, and the potential for artificial intelligence (AI) to transform the eDiscovery process for legal professionals.

Over the past decade, agencies have seen a significant increase in audio and video-based evidence in legal proceedings. With the proliferation of cell phones, body-worn cameras, Zoom calls, CCTV systems and more, the majority of evidence in many cases is video-based. Up until a few years ago, a significant amount of human labor was required to manually process vast amounts of audio and visual data during investigations and the legal discovery process. With advances in technology, including AI, reviewing and analyzing this kind of data is proving more efficient and cost-effective.

Participants agree that while the use of AI technologies is emerging, the potential to improve processes, increase privacy, and create efficiencies with AI is considerable. However, while some advances in technology have been made, most agencies are not currently equipped to adequately address the anticipated increase in AV data evidence in the coming years.

#### Challenges with AV Evidence Processing

Roundtable participants from different federal agencies experience similar challenges and limitations with current eDiscovery processes and technology. Federal agencies charged with investigating and prosecuting crimes can receive thousands of tips from the public in the form of multi-media content such as personal video, images, voicemails, or social media. Depending on the nature and scale of the investigation, agencies may also review data from body worn cameras and close circuit TV.

In certain cases, investigative and prosecutorial organizations must have the ability to not only index and catalogue myriad data sets, but also standardize the data in order to share it

with others, including state and local partners and outside legal defense teams. Currently, investigative and prosecutorial agencies receive data in different formats, making standardizing and storing this data more difficult. Converting data into reliable formats that can be used in court can also prove challenging.

#### Too Much Data???

Investigative organizations lack ability to:

- ❖ Index and catalogue myriad data sets
- ❖ Standardize the data to share with partners
- ❖ Convert data into reliable formats for better storage
- ❖ Provide sufficient employee training
- ❖ Collaborate between data analysts and data litigators due to the incompatible tools used by each

Another challenge most roundtable participants experience is the lack of proper training on the various systems used to share eDiscovery data with agency personnel and outside defense attorneys. Due to the lack of consistent data sets and programs, agencies continually train defense attorneys from smaller firms on how to access discoverable data. Moreover, most of the tools used to forensically analyze the data are not compatible with the tools used by litigators. There remains a significant human element to adequately processing eDiscovery data due to technology limitations. While there are some limitations, agencies have experienced significant advancements within the past several years.

#### Advancements in eDiscovery

As AI becomes more commonplace and understood within the federal government, more agencies are seeing the benefits of introducing AI to automate processes and create workflow efficiencies. With AI, agencies can create rules and run models on multiple, disparate data sets at scale, creating efficiencies, saving resources, and improving privacy.

Some agencies are looking to use robotics processing automation (RPA) as a tool to standardize data and create repeatable processes by bridging disparate data sets. While it is currently a politically charged concept, facial recognition is another beneficial AI tool that serves multiple purposes and can create efficiencies while protecting personal information. Facial recognition can not only help investigators track objects or identify individuals perpetuating crime across multiple media platforms, but can also help to redact private and sensitive information from videos.

AI automation and tagging can also help with the redaction of unnecessary information, testimony, or detail while maintaining need-to-know information and protecting the privacy of other individuals involved. The use of AI to identify information for redaction is particularly helpful in cases involving child pornography.

A provision of providing government service to the public often requires a certain level of authentication, such as when receiving state or federal licenses, benefits or identification cards. Currently, personal identities are most often authenticated by person. With the introduction of AI and facial recognition technology, authentication of personal identity can be conducted much more quickly and efficiently.

Roundtable participants stressed the importance of understanding that applicable personnel are trained in the best practices and standards of facial recognition use. Those who use this form of AI technology understand that the output of facial recognition is never a basis for identification on its own. Information gathered from facial recognition should be considered like any other quality tip and must undergo similar due diligence before taking action on the intelligence. Not only can facial recognition serve as a method of identifying potential defendants, it can also help defendants exculpate themselves if they were to be in a different location at the time of the alleged crime.

While the use of facial recognition AI technology is proving useful and efficient, agencies are facing increased threats from the proliferation of synthetic media, or the manipulation of media, and the use of deepfake identities. While most media is manipulated to some extent, agencies consider the intent behind the manipulation when analyzing data in legal proceedings. While deepfake technology is relatively primitive, fraud and sextortion are concerning consequences

of this form of impersonation. Investigators and prosecutors are often challenged with the ability to prosecute certain instances where there is no true victim.

Federal agencies routinely partner with private entities to share information and collaborate on solutions to address and prevent the spread of synthetic media on social networks. Private sector, including social media companies, are taking the lead to develop open standards for multi-level authenticity in an effort to prove the authenticity of content made on their platforms. The ability to trace any edits made to content through a chain of authenticity will help identify synthetic media and potentially stop its viral spread.

## Future of eDiscovery

Roundtable participants attribute the current level of technological advancement and cross departmental collaboration to the recognition from leadership that this technology should be made available across entire departments, not just specialized groups. Many smaller departments and agencies do not have budgets to invest in such technological advancements, so collaboration and sharing of tools among agencies is critical to overcome current and future challenges.

Future policies and legal guidance from the government on the use of AI technologies should consider the civil rights, liberties, and privacy of individuals, and provide guidance on how AI is defined in certain contexts. This is crucial as technology continues to evolve and become more ubiquitous.

In addition to partnering with the private sector to solve problems, the government should be transparent with its needs and challenges to private developers to avoid transcoding issues. Future tools must work with existing technology to ensure ease of use and minimal training.

Advancements in technology are inevitable, so roundtable participants agree that considerable focus should be placed on training and retaining personnel on this emerging technology, especially in the legal space. A skilled workforce is critical to the success of any AI integration and automation of the eDiscovery process.

[Contact us](#) today to learn more and get involved in the Artificial Intelligence Working Group!